Exam **Spatial Statistics**,
June 29, 2023, 1.45–4.45 pm.

This is a closed book exam. The use of electronic devices is not allowed. Please answer all questions clearly and legibly. For each of the five questions, a correct answer is worth ten points. The final mark is the sum of all points divided by fifty and rounded to an integer. Please make sure that your name and student identification number are on every sheet of paper that you hand in.

---

1. Dutch regulations for archeological conservation require that the depth to a particular, undisturbed soil layer is investigated in agricultural parcels. The depth to this layer should be within 30 cm of the soil surface, but it may be varying due to spatial variability. Such an undisturbed layer may contain valuable archaeological remnants.

   For each municipality M, one may have some $N = 150$ agricultural parcels, again a number that varies per municipality. For each municipality, a fixed number $n = 10$ parcels have to be selected, and on each parcel $k = 4$ samples are to be taken to determine the depth to the undisturbed layer.

   a Describe how design-based sampling should take place to investigate the depth to the undisturbed layer.

   b Alternatively, model-based sampling strategy could be applied. An objective (fitness) function could be that on average the depth to the undisturbed layer is best mapped. Describe how model based sampling could be executed. You may assume that a variogram is available.

   c Give briefly you the pros and cons of the two methods.

2. **a** Consider the second-order and intrinsic stationarities. Which is simpler to estimate the spatial dependency and why?

   b Consider the ordinary kriging variance

$$\sigma^2(s_0) = \sigma^2 - c_0^T C^{-1} c_0 + (1 - c_0^T C^{-1} 1)(1^T C^{-1} 1)^{-1}(1 - c_0^T C^{-1} 1)^T$$

   where σ^2 is the variance of the data, c_0 is the covariances between the prediction location and the observation locations, and C is the covariances between the observations.

   i Explain the role of the terms last term on the right hand side of the equation.

   ii Explain why the kriging variance dependent on the data configuration and independent of data values.

   c Consider the exponential semivariance model below where

$$\gamma(h) = 2.5\left\{1 - \exp\left(\frac{-h}{30}\right)\right\}$$

   where h is the lag distance in kilometers (km). If the effective range, usually taken as the distance at which γ equals 95% of the sill is 90 km

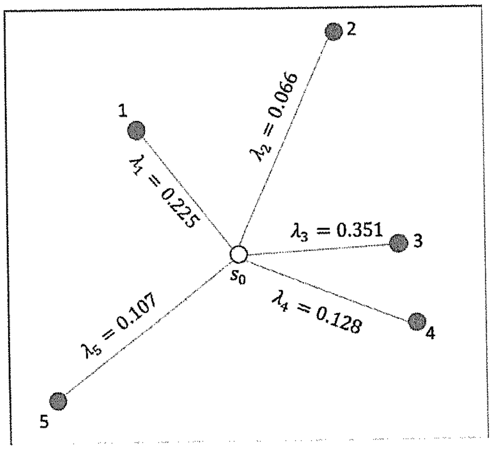
   i Write a expression for the corresponding covariance model.

   ii What is the sill

   iii What is the semivariance between observations separated by 100 km

p.t.o.

3.



a Consider the configuration of five sites shown in the figure above. Using the variogram information provided in **2c** above, the kriging weights λ_i, $i = 1, \ldots, 5$, were estimated for each observation location based on the assumption of known constant mean of $\mu = 160.37$. Estimate the value at s_0 using the observed data values

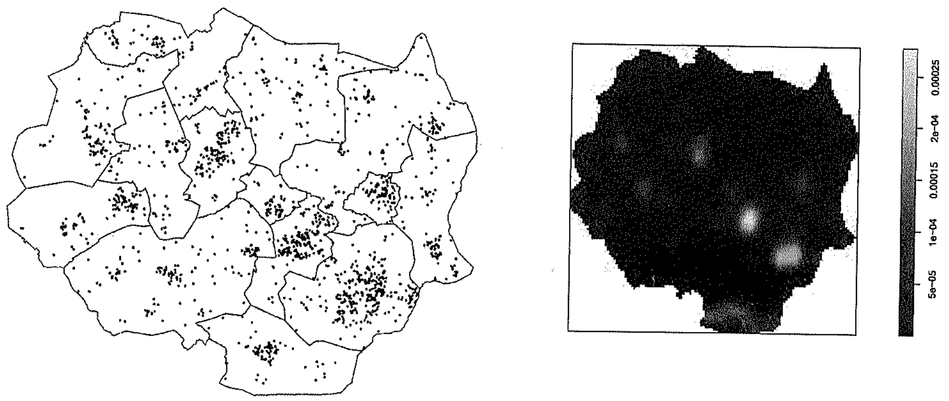
Sites	y
1	122.0
2	183.0
3	148.0
4	160.0
5	176.0

b The estimated value at s_0 is considered as the Best Linear Unbiased Predictor (BLUP). Briefly comment on why?

p.t.o.

4. **a** Give the definition of the intensity function of a point process and a formula for its kernel estimator. Is the estimator unbiased?

   b Discuss some alternative approaches for estimating the intensity function and elaborate on their relative advantages and disadvantages.



   c Consider the figure above. Its left panel displays a mapped point pattern of chimney fires that occurred in the Twente region during 2004–2020. In the right panel, a density map of buildings susceptible to chimney fires (in number per square km) is shown. Please give a qualitative description of the two plots and the relationship between them, if any.

   d Suggest a plausible model assuming that there are no interactions between the points. How would you validate the fitted model?

5. Let X be a Poisson process on $[0,10]^2$ with intensity function

$$\lambda(x,y) = \alpha x + \beta y, \quad x, y \in [0,10]$$

and unknown parameters $\alpha, \beta \in \mathbb{R}$.

   a For what values of α and β is λ a valid intensity function?

   b Calculate the void probability $v(A)$ for

      i $A = [0,1]^2$, and

      ii $A = [9,10]^2$.

   c Derive the joint probability density function of the random vector $(N_X([0,1]^2), N_X([9,10]^2)$, writing $N_X(A)$ for the number of points in $X \cap A$, $A \subset [0,10]^2$.

   d Suppose that $\alpha = \beta$ so that $\lambda(x,y) = \alpha(x+y)$. Does a maximum likelihood estimator for the unknown parameter α exist? If so, give a closed form expression.

   e For the case that α and β may be different and are both unknown, does a maximum likelihood estimator for the unknown parameter vector (α, β) exist? If so, derive the estimator.